Comparative analysis of transcriptome and fitness profiles reveals general and condition-specific cellular functions involved in adaptation to environmental change in Saccharomyces cerevisiae


DOI
10.1089/omi.2010.0049

Publication date
2010

Document Version
Final published version

Published in
Oomics

Link to publication

Citation for published version (APA):
Comparative Analysis of Transcriptome and Fitness Profiles Reveals General and Condition-Specific Cellular Functions Involved in Adaptation to Environmental Change in *Saccharomyces cerevisiae*

Anna Zakrzewska,1,* Andre Boorsma,2,† Alex Ter Beek,1 Jos A. Hageman,3,‡ Johan A. Westerhuis,3 Klaas J. Hellingwerf,3 Stanley Brul,1 Frans M. Klis,1 and Gertien J. Smits1

Abstract

The transcriptional responses of yeast cells to a wide variety of stress conditions have been studied extensively. In addition, deletion mutant collections have been widely used to measure the combined effect of gene loss and stress on growth (fitness). Here we present a comparative analysis of 1,095 publicly available transcription and genome-wide fitness profiles in yeast, from different laboratories and experimental platforms. We analyzed these data, using T-profileer, to describe the correlation in behavior of *a priori* defined functional groups. Two-mode clustering analysis of the fitness T-profiles revealed that functional groups involved in regulating ribosome biogenesis and translation offer general stress resistance. These groups are closely related to growth rate and nutrient availability. General stress sensitivity was found in deletion mutant groups functioning in intracellular vesicular transport, actin cytoskeleton organization, and cell polarity, indicating that they play an key role in maintaining yeast adaptability. Analysis of the phenotypic and transcriptional variability of our *a priori* defined functional groups showed that the quantitative effect on fitness of both resistant and sensitive groups is highly condition-dependent. Finally, we discuss the implications of our results for combinatorial drug design.

Introduction

Understanding the complexity of cellular responses to changes in the external environment is still one of the major challenges in biology. Genomic tools, which probe the complexity of cellular functioning, such as global transcription or fitness measurements, were first developed for the model organism *Saccharomyces cerevisiae*. Global transcription experiments measure in a single assay the changes in expression levels of all genes in response to altered environmental conditions and as such provide snapshots of the changes occurring in genome-wide expression programs in response to various stresses. For example, the study of stress-induced global expression levels in *S. cerevisiae* led to the discovery of the environmental stress response program (ESR). This program, which is activated in response to many stress conditions, is characterized by the upregulation of gene expression mediated by the Msn2p and Msn4p transcription factors and downregulation of the expression of genes involved in ribosome biogenesis (Causton et al., 2001; Gasch and Werner-Washburne, 2002; Gasch et al., 2000).

In recent years, genomic collections of yeast deletion mutants have been exposed to many different conditions, greatly expanding the data on genome-wide behavior. The types of experiments and platforms vary greatly, which has hindered comparative analysis on the level of individual genes. Additionally, several studies of global fitness and transcript levels reported that individual genes required for optimal growth are often not activated transcriptionally under the same conditions (Birrell et al., 2002; Giaever et al., 2002; Smith et al., 2006; Tai et al., 2007). These studies indicate that our understanding of cellular functions required for coping with imposed stresses and their regulation in response to environmental changes is far from complete. To address this issue...
we performed a comparative analysis of genome-wide studies on the effect of single gene deletions on growth, and integrated this with a similar analysis on the level of transcription, in order to dissect the global changes occurring on different levels of biological control. Using T-profiler, a method based on t-statistics, we directly analyzed the behavior of a priori defined groups, from 159 fitness and 936 transcription experiments (Boorsma et al., 2005, 2008).

We confirm that transcription factor-controlled gene expression is not propagated onto the phenotypic level, with the exception of ribosomal protein-encoding genes, which form a tightly coregulated functional group controlled by a small set of transcription factors. Although similar numbers of significantly altered functional groups are identified in both transcription and fitness profiles, they represent very distinct functions. Two-mode clustering of t-values of Gene Ontology (GO)-based groups and conditions for all fitness t-profiles reveals strong interactions between functional groups and treatments. Reduced functioning of ribosomal biogenesis and translation leads to increased fitness in many, but not all, conditions. In addition, there is a general stress sensitivity of deletion mutants involved in the intracellular vesicular transport machinery and the actin cytoskeleton, functions that seem crucial in the general adaptive mechanisms of S. cerevisiae. Analysis of the variability of a group’s transcriptional responses and of the effect they have on fitness throughout the compendium suggests that the regulation of stress-sensitive groups is mostly posttranscriptional. Additionally, the quantitative effect of perturbation of a specific function is highly condition-specific, which provides targets for combinatorial design of antifungal treatment. In conclusion, we report global properties in yeast fitness data and provide a quantitative comparison of global fitness and transcription differences on the level of cellular functions and biological processes. Our results provide a better understanding of yeast adaptation to many conditions, and of the cellular functions required for this adaptation.

Materials and Methods

Libraries of genome-wide fitness and transcription data

Our library of fitness data contains data of 159 fitness experiments carried out with all nonessential deletion mutants of S. cerevisiae (Table 1). The yeast deletion collection can be tested for susceptibility to any type of stress in at least three various ways. The growth of individual mutants can be detected, the colony size of each mutant can be determined, and finally, by amplification of unique tags for each mutant, the abundance of all mutants can be scored simultaneously on a microarray. We used datasets exploring both haploid and homozygous diploid yeast knockout strains, yielding a broad range of datasets that are not easily comparable. Our fitness library contains data from all three different experimental platforms and includes experiments where cells are subjected to various stress conditions. The library of gene transcription data contains data of 936 hybridization experiments carried out with S. cerevisiae from 19 publications (Boorsma et al., 2008). This transcription library contains data from different DNA-array platforms such as GeneFilter, Affymetrix, and spotted slides, and includes experiments with gene deletion strains, synchronized cells for cell cycle analysis, sporulating cells, and cells subjected to various physical and chemical perturbations.

T-profiler

We use T-profiler (www-t-profiler.org) to calculate the t-values corresponding to the average fitness levels of predefined groups of deletion mutants (mutant groups) and expression levels of predefined gene groups. Previously, the same method was used to quantify the average expression levels of predefined gene groups in genomic expression and fitness experiments (Boorsma et al., 2005; Zakrzewska et al., 2005, 2007). We use the same method, except for the fact that gene groups now become mutant groups (see below for a description of each category of mutant groups), and their t-scores are calculated with respect to the total population of viable haploid mutants sampled. Both gene and mutant groups can contain the same members, yet to avoid confusion we generally use the term “gene group” for the description of transcription data, and the term “mutant group” for the description of fitness data. The fitness and transcriptome t-profile datasets in our library have been analyzed using T-profiler and the resulting fitness and transcription t-profiles have been uploaded to two databases, which can be found at http://www.science.uva.nl/~boorsma/T-base-fitness/ and http://www.science.uva.nl/~boorsma/T-base-all, respectively.

<table>
<thead>
<tr>
<th>Literature</th>
<th>Deletion pool used</th>
<th>Type of platform</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Birrell et al., 2001)</td>
<td>Homozygous nonessential diploid</td>
<td>Chip</td>
<td>2</td>
</tr>
<tr>
<td>(Birrell et al., 2002)</td>
<td>Homozygous nonessential diploid</td>
<td>Chip</td>
<td>4</td>
</tr>
<tr>
<td>(Brown et al., 2006)</td>
<td>Homozygous nonessential diploid</td>
<td>Chip</td>
<td>27</td>
</tr>
<tr>
<td>(Dudley et al., 2005)</td>
<td>Homozygous nonessential diploid</td>
<td>Colony</td>
<td>21</td>
</tr>
<tr>
<td>(Giaever et al., 2002)</td>
<td>Homozygous nonessential</td>
<td>Chip</td>
<td>6</td>
</tr>
<tr>
<td>(Parsons et al., 2006)</td>
<td>Haploid nonessential</td>
<td>Chip</td>
<td>84</td>
</tr>
<tr>
<td>(Warringer et al., 2003)</td>
<td>Haploid nonessential</td>
<td>Individual</td>
<td>4</td>
</tr>
<tr>
<td>(Wu et al., 2004)</td>
<td>Homozygous nonessential diploid</td>
<td>Chip</td>
<td>8</td>
</tr>
<tr>
<td>(Zakrzewska et al., 2007)</td>
<td>Homozygous nonessential diploid</td>
<td>Chip</td>
<td>2</td>
</tr>
</tbody>
</table>

The list contains publication from which experimental data was obtained, the type of yeast deletion pool used, the platform that was used to perform the genome-wide experiment, and the number of experiments per study.
Transcription factor (TF) regulated groups

We specify TF-regulated groups in two ways; Motif-based groups are made up of those genes that contain a common regulatory motif in the 600 bp of their promoter region allowing no overlap between neighboring ORFs (van Helden et al., 2000). The consensus motifs used in T-profiler (Boorsma et al., 2005) are derived from three different sources. First, motifs were extracted from the SCPD database (http://rulai.cshl.edu/SCPD/). Additionally, motifs were found by comparing the genome sequence of highly related yeast species (Gasch et al., 2000; Kellis et al., 2003). Finally, motifs were discovered in various microarray experiments by the REDUCE algorithm (Bussemaker et al., 2001; Roven and Bussemaker, 2003). Most of these motifs are similar or identical to motifs described in the literature. In total, 153 motif groups have been included in T-profiler calculations. Transcription factor occupancy (TFO)-based groups were generated using the TFO data from chromatin immunoprecipitation microarray (ChIP-chip) studies as input in T-profiler (Harbison et al., 2004; Lee et al., 2002). These datasets contain results of ChIP-chip experiments performed in rich medium (YPD) with 203 transcription factors. For 84 of these binding to promoter regions was also measured in at least 1 of 12 other environmental conditions. A mutant was considered to be part of a TFO group if the p-value reported for its gene by the authors was smaller than 0.001. In addition, TFO groups were required to have at least seven gene members. This resulted in 252 TFO groups that were used for T-profiler analysis.

Functional groups

GO-based groups contain the genes/mutants associated with a specific GO category as well as all of its child categories. Only GO groups with at least seven members were used for calculation. This approach resulted in a reduction of the original Saccharomyces Genome Database (SGD) to 3,836 GO-based groups to 1,346 GO-based groups, which were used for T-profiler analysis. Significantly scoring GO-based mutant groups directly indicate which functions or cellular processes are required for optimal growth under a specified condition.

MIPS-based groups contain genes annotated by the MIPS organization as well as three types of categories that were analyzed together with MIPS-derived groups. Three major categories of groups are based on functional classification, protein complex and cellular localization. We only used the MIPS mutant groups with at least seven members for our analysis. Significantly scoring MIPS-based mutant groups directly indicate which functions are required to maintain optimal fitness in a specific stress condition. Three minor categories were added, based on the morphology database, the phenotype database, and the synthetic lethality database. For the first, 4,400 nonessential mutants were scored for their cellular shape and subsequently annotated to 11 categories. We used six of these morphology groups, based on the number of their members and excluding the morphology group named WT (wild type). We tested 233 phenotype categories derived from the MIPS database that describe various features of mutant strains ranging from mutants’ drug susceptibilities, defects in cell shape or organelle morphology, cell cycle, mating, sporulation, to auxotrophies, and defects in utilization of carbon and nitrogen. The weakness of this database is such that it is not complete, due to the fact that it is only a collection of available facts from literature. However, because it is entirely based on analysis of mutants, it is very suitable for characterization of fitness data. The Synthetic Lethality Interactions database comprises 462 groups of mutants, representing a full spectrum of biological processes and cellular functions. The data contained there comes from genetic screens where a deletion mutant of a query gene is crossed into an array of viable gene deletions, which generates an array of double mutants that can be scored for synthetic lethal or sick (reduced fitness) phenotypes. The result of such screens provides information on genes that are involved in the same function or genes, which are able to buffer each other’s functions (Tong et al., 2004). This database is not yet complete, as only around 10% of genes were tested as a query in the synthetic genetic interaction network. However, the enormous advantage of the Synthetic Lethality categories is that they are determined by the organism and not annotated by a researcher.

Two-mode clustering of fitness t-profiles

The t-values of all GO-groups in all experiments make up a two-dimensional data matrix. To find discreet clusters of both GO-groups and experiments that are most similar, we used two-mode clustering (Van Mechelen et al., 2004). We used a Genetic Algorithm (GA) based two-mode clustering method, that searches simultaneously for the optimal clustering of groups and experiments by minimizing the summed and squared deviation of all t-values from the mean within one cluster (Hageman et al., 2008; Wehrens and Buydens, 1998). In this way, groups that behave similarly under certain experimental conditions will be clustered together. The Matlab routines of GA-based two mode clustering routines applied in this article are available at http://www.bdagroup.nl (Hageman et al., 2008). A GA requires a predefined number of clusters on both axes. We chose the knee method to decide on the correct number of clusters. This method finds the knee or “L” in a contour plot of the number of clusters versus the sum of squares of the residuals (Salvador and Chan, 2004). The knee is a combination of cluster numbers for which an additional cluster no longer sharply decreases the sum of squares of the residuals. No unique L could be recognized, and we decided on 10 GO ontology group clusters, and nine experiment clusters, because increasing these numbers did not lead to a strong increase in performance.

Analysis of group variability

To systematically analyze group behavior over all datasets, we calculated for each group its general position on a scale of resistant to sensitive, by calculating the t-value for all its t-scores in the 159 fitness experiments compared to the t-scores of all other groups (Tmean), as well as its position on a scale from minimally modified to highly modified, by calculating a t-value for the deviations of the group’s t-scores from the average t-score over all experiments (Tvar). Similarly, we calculated these values for all groups in all 936 transcriptome datasets to understand transcriptional variability of groups.
from all individual mutants into a set of \( t \)-values for all mutant groups, resulting in a fitness \( t \)-profile (Fig. 1). Similarly, a transcription \( t \)-profile is the collection of all \( t \)-values of pre-defined gene groups for each individual transcription experiment in our collection of 936 such experiments (see Boorsma et al., 2008). One of the important advantages of T-profiler over other methods used to analyze large-scale datasets is that the transformation of the data into a number of \( t \)-values for all groups introduces an intrinsic normalization. It therefore allows for direct comparison of datasets from various types of experiments, including colony size measurements, arrays for assessing abundance of all deletion mutants by unique DNA tags, individual growth curve monitoring or the use of haploid or diploid strains. Importantly, it also allows comparison of fitness and transcription experiments (Boorsma et al., 2008; Zakrzewska et al., 2005, 2007). Second, the group-wise analysis gives greater robustness against possible errors in the determination of fitness of individual mutants, both because of the group size (Kim and Volsky, 2005) and because in T-profiler the single highest and lowest value in the group are discarded, thus minimizing the effect of outliers (Boorsma et al., 2005). Last, T-profiler is capable of detecting minor changes in average fitness, which means that individual mutants do not need to show significant changes in fitness in order for the group to show a significant change, if the changes are strongly correlated. This property of T-profiler analysis provides a more realistic insight into cellular responses than any stringent cutoff, as many biological processes tend to be regulated in subtle and gradual ways (Bussemaker et al., 2007).

**Functionally related genes are required for optimal growth under stress conditions**

To respond and adapt to varying growth conditions particular functions and cellular processes need to be activated or repressed upon change. We speculated that mutation of genes that participate in adaptation to an environment should result in the mutants sharing similar phenotypes or fitness properties. We found that yeast strains deleted for genes regulated by the same transcription factors do not generally share similar fitness properties. Although groups of genes controlled by specific transcription factors did show coordinated behavior in transcription datasets, they did not appear as significant groups in fitness datasets (Fig. 2a and b). The exception to this rule were the groups controlled by Fhl1p, Rap1p, and Sfp1p (Supplementary Table 1). All three transcription factors are almost exclusively involved in regulating the transcription of ribosomal proteins and genes involved in ribosome biogenesis (Zhao et al., 2006). The corresponding gene and deletion mutant groups were significant in many transcriptional datasets, but also in up to 50% of the fitness datasets.

To examine whether fitness properties of deletion mutants are linked to functional characteristics other than transcription factor regulated gene expression, we studied the distribution of gene and deletion mutant groups derived from GO and MIPS in both transcription and fitness and datasets (Fig. 2c and d). The majority of fitness \( t \)-profiles (81%) and transcription \( t \)-profiles (61%) scored more than 10 significantly changed GO-based groups. Similarly, above 80% of all fitness and transcription \( t \)-profiles scored more than 10 significantly changed MIPS-based groups. This shows that many func-

\[
\begin{align*}
\text{t}_{\text{mean}} &= \frac{\mu_t - \mu_t'}{S_t \sqrt{\frac{1}{N} - \frac{1}{N'}}} \\
\text{with the pooled standard deviation}
\end{align*}
\]

\[
S_t = \sqrt{\frac{(N_t - 1) \times S_t^2 + (N_t' - 1) \times S_t'^2}{N_t - N_t' - 2}}
\]

Here \( \mu_t \) indicates the average \( t \)-value and \( s_t \) the standard deviation of one group over \( N_t \) experiments, whereas \( \mu_t' \) is the average \( t \)-value and \( s_t' \) the standard deviation for all other groups over all experiments, adding up to \( N_t' \). For the variability the absolute deviation of the group \( t \)-value in each experiment from the average \( t \)-value of this group over 159 or 936 experiments was calculated. A \( t \)-value was calculated (\( T_{\text{var}} \)), similar to that for the mean, but now using the set of deviations instead of actual \( t \)-values. Associated \( p \)-values were corrected for the total number of \( t \)-values in the analysis, and a corrected \( p \)-value of 0.05 or lower was considered significant.

**Results**

**Generation and analysis of fitness \( t \)-profiles**

We collected 159 global fitness datasets of homozygous diploid or haploid deletion strains of all nonessential genes of *S. cerevisiae* exposed to various environmental conditions, including two datasets of our own (Table 1). Each fitness experiment provides a quantitative description of a whole spectrum of susceptibilities, ranging from highly sensitive to highly resistant, of all individual deletion mutants, represented by the log2 ratio of abundance of the single mutants in control versus treatment. We reanalyzed the fitness of mutant groups compared to the whole population using T-profiler (Boorsma et al., 2005), which converts fitness values into functionally interpretable groups. (A) Fitness phenotypes of mutant collections can be represented as a distribution. Within this distribution, a *priori* defined groups can be significantly sensitive (yellow dots) or resistant (blue dots) as a group, even when the individual members of the group may not all have significant phenotypes. (B) T-profiler converts fitness or transcriptome datasets into fitness of transcription \( t \)-profiles, which implies an inherent normalization and yields an easily interpretable functional representation of the data.
Functional groups are transcriptionally regulated, and deletion mutants in functional groups often have correlated growth effects. However, we cannot determine from these results whether the functional groups in both types of data are the same. To further investigate this point we performed a direct analysis of the individual groups that were transcriptionally regulated or those that resulted in similar phenotypes, in a subset of experiments with identical treatment times and dosages of three DNA-damaging agents (Birrell et al., 2001). Our data revealed that there was no correlation of GO groups between the fitness and transcription t-profiles (Supplementary Fig. 1).

These results show (Fig. 2) that genes required for optimal growth under specific conditions generally function together in biological processes, or that their gene products localize to the same cellular compartments and/or form complexes, or that their deletion mutants share morphological characteristics. However, the genes in these groups are rarely transcriptionally coregulated, while transcriptionally coregulated functional groups do not share fitness characteristics.

**Two-mode clustering of fitness t-profiles reveals general fitness characteristics of *Saccharomyces cerevisiae***

We next asked whether the fitness data reveal general features required for adaptation to suboptimal growth conditions, similar to the well-described ESR in transcriptome
experiments (Gasch et al., 2000; Gasch and Werner-Washburne, 2002). Above, we showed that yeast fitness is not defined by transcriptional responses, but that mutations in genes that belong to specific functional groups often lead to similar effects on fitness. To analyze behavioral correlation between different functional groups, we performed two-mode clustering analysis (Hageman et al., 2008) of the 561 GO groups with significant fitness $t$-scores in at least 4 out of the 159 experiments. Two-mode clustering minimizes deviations within whole clusters, rendering this method exceptionally suitable to reveal general traits and to emphasize interactions between functions and growth conditions. We determined an optimal combination of 10 GO group clusters (rows, a to j), and 9 condition clusters (columns, i to ix). The results of the clustering are presented in Figure 3, where yellow represents sensitivity, and blue represents resistance of the GO-defined mutant groups in fitness experiments. The full lists of groups belonging to each cluster can be found in Supplementary Tables 2 and 3, and the full

FIG. 3. Two-mode cluster analysis of GO-based fitness $t$-profiles. The $t$-values in all 159 fitness experiments for the GO-based mutant groups that were significant in at least four fitness experiments were clustered into 10 group clusters and 9 experimental clusters using a GA-based two-mode clustering algorithm (See Materials and Methods). Each column represents a single fitness $t$-profile, and each row represents a single GO-based group. Blue indicates resistant, and yellow indicates sensitive groups. For a complete description of all clusters see Supplementary data.
clustering is available in the Supplementary materials. Experiments from different laboratories analyzing the effect of caffeine (Brown et al., 2006; Parsons et al., 2006), UV (Birrell et al., 2001; Dudley et al., 2005), cisplatin (Parsons et al., 2006; Wu et al., 2004), hydroxyurea (Brown et al., 2006; Dudley et al., 2005), and hygromycin B (Brown et al., 2006; Dudley et al., 2005; Parsons et al., 2006) clustered together, which shows that our method allows for accurate comparison of data from different laboratories, different ways of screening and analysis, as well as different dosages of the treatment. The interaction between functional groups and experiments is exemplified in cluster d: the groups belonging to this cluster are generally mildly stress resistant, except when they are exposed to the stress conditions in cluster vi. Interestingly, all functional groups in d are closely related to DNA damage repair and cell cycle regulation, and all stress conditions in vi are known to cause DNA damage. This reveals that tolerance toward DNA damage depends on a specific set of cellular functions, that are not required for adaptation to other conditions.

Several clusters appeared generally stress resistant, meaning that the average stress-induced growth reduction of the deletion mutants in these groups is significantly less than the average growth rate reduction of all other deletion mutants in the population. In clusters a, b, and c we find generally resistant groups, with the most strongly resistant groups in cluster a. Most of the groups in this cluster are themed around the cytosolic ribosomes and protein synthesis. In clusters b and c we find mostly groups involved in mitochondria and respiratory energy metabolism, as well as groups involved in RNA splicing and ribonucleotide metabolism, respectively. Although the groups in clusters a and b are resistant to most of the environmental changes, they show a relatively strong growth reduction in treatment cluster vii, and also to a lesser extent in viii and ix. Cluster vii contains all the conditions using rich media with glucose or alternative carbon sources. Other conditions in vii, viii, and ix in which the groups in a and b have sensitive growth phenotypes include synthetic complete media, minimal media, amino acid limited growth, as well as rapamycin- and sodium azide-induced stress. This suggests that the groups in clusters a and b are sensitive under conditions where nutrient availability limits growth, or in which nutrient signaling or nutrient-dependent regulation is targeted, although they are resistant under conditions that can be considered as nonphysiological stress. The groups in these three clusters therefore seem to be required for nutrient and energy regulation of growth. Interestingly, the groups in cluster c do not show this sensitivity in growth related conditions. These groups are slow growers onYPD, but are on average neutral to resistant in clusters vii, viii, and ix. Although the groups in clusters a, b, and c strongly resemble the groups downregulated in the transcriptionally apparent ESR, the groups in RNA splicing and ribonucleotide metabolism seem to separate on a phenotypic level from those involved in protein synthesis and energy generation. Additionally, the groups in the mitochondrial and respiratory cluster b show a rare sensitivity for several stresses in cluster ii, in addition to those in clusters vii, viii, and ix. They are shared only with three groups from cluster a, groups that are exceptional in this cluster because they are the only mitochondria-related groups.

Clusters e, f, and g are less well defined, because the groups in these clusters are significant only in some experiments. However, the groups in cluster f have resistances and sensitivities similar to those in a and b, albeit less pronounced. Indeed, many functional groups in this cluster are involved in energy-generating metabolism and mitochondrial function. Also, in this cluster we find several groups involved in chromatin remodeling (a.o. SWI/SNF complex) and histone acetylation (a.o. SAGA complex, histone acetylation), whereas the groups involved in histone deacetylation localize to cluster g, together with many growth-related (a.o. cell regulation, growth of cell cycle) and polarity-related functional groups (a.o. site of polarized growth, Rho protein signal transduction, organelle localization). The most remarkable difference in the sensitivities of the groups in clusters f and g lies in the resistance shown by the latter groups in experimental clusters vii and ix, with slow growth and amino acid limitations, respectively.

Remarkably, clusters h, i and j contain mutant groups that show sensitivity in the vast majority of the experiments compared to the accompanying control experiment. These functions apparently help in withstanding environmental changes in general. In cluster j we find groups that are especially important for growth under suboptimal conditions, and almost all of these are involved in vesicle-mediated transport, centered around the Golgi, the endosome, and the vacuole. In clusters h and i we also find categories involved in vesicular transport, as well as several functions related to vesicles and vesicular transport, such as lipid synthesis, actin cytoskeletal function, and protein localization. Sensitivity of the groups in this cluster is low when cells are exposed to treatments in cluster ix, including for instance DTT, arsenic, and phenanthrolone, suggesting that these treatments cannot further reduce growth in mutants in which secretory functions are already defective. Interestingly, sensitivity is highest under conditions in clusters iii and v, with many treatments affecting signal transduction and regulation of cell polarity, such as caffeine, wortmannin, FK506, latrunculin, tunicamycin, etc. Clearly, cellular remodeling upon changing environmental conditions requires massive rearrangements and adaptation of membrane and organelle protein content.

**Group fitness dynamics separate functions important for general cellular maintenance from those important for adaptability**

As observed in Figure 3, functional groups can be either generally sensitive or resistant, or largely stress specific. Also, group fitness can either vary widely, or show very little variation. We hypothesized that groups with low fitness dynamics are either generally important, if they are always sensitive, or are highly condition specific, if they tend to be neutral to resistant. In contrast, highly dynamic groups should be important for adaptability, as their role varies strongly with the changing environment. To test our hypothesis, for each group we calculated its general position on a scale of resistant to sensitive ($T_{mean}$), as well as its position on a scale from minimally dynamic to highly dynamic ($T_{var}$) (see Materials and Methods), compared to all other groups. This allows us to systematically analyze the variability of functional group behavior, and to determine for all groups whether they are important for general cellular maintenance, or for adaptability.

A plot of $T_{var}$ versus $T_{mean}$ is depicted in Figure 4A, and was colored depending on the clustering shown in Figure 3.
fitness experiments are plotted versus p.<br>

Wide dots represent those with significant (corrected to these groups have very little effect on fitness in our selected only groups with significant are absent in the two-mode cluster analysis, because we see-<br>
data) in the lower part of Figure 4A. Most of these groups<br>

FIG. 4. Variability of GO-ontology based functional groups in fitness and gene expression data. (A) The t-values for fitness variability (Tv) of GO-based groups in all 159 global fitness experiments are plotted versus t-values for mean fitness (Tmean) of the same groups in all experiments. Wide dots represent those with significant (corrected p-value <0.05) Tvar fitness. (B) The t-values for transcriptional variability of GO-based groups in all 936 transcriptome experiments are plotted versus the Tvar of the same groups in the fitness data. Wide dots represent those with significant (corrected p-value <0.05) Tvar transcription. Colors correspond to the clusters generated in Figure 3 (see legend). For variability data of all GO groups see Supplementary data.<br>

We observed a strong aggregation of groups in the generally stress resistant clusters a, b, and c (see Fig. 3) in the upper left, and groups belonging to the generally sensitive clusters h, i, and j in the upper right corner of the graph. This confirms that functions, which strongly determine yeast fitness upon environmental changes, are generally highly dynamic.<br>

We find many functional groups that have significantly low variability (438 out of 1,186 tested, see Supplementary data) in the lower part of Figure 4A. Most of these groups are absent in the two-mode cluster analysis, because we selected only groups with significant t-scores in four or more experiments for the clustering. Mutations in genes belonging to these groups have very little effect on fitness in our set of experiments, either positively or negatively. It is possible that the corresponding functions do not actually have low variability in all conditions, but are not challenged in the experiments analyzed, for example, because they are important for specific developmental pathways such as pseudohyphal growth, mating, or sporulation. Alternatively, the fitness of deletion mutants in these groups does not vary much because the effect of deletion of genes in these functions is buffered. Indeed, GO-ontology groups such as karyogamy and conjugation with cellular fusion, spore wall and pro-spore membrane, pseudohyphal growth, but also replicative cell aging and nicotinamide metabolic process, are among the significant low variability groups (see Supplementary data), and are also significantly resistant, suggesting that the presence of these specific functions leads to nonoptimal growth in most cases, and that the inactivation of these functions causes the cells to grow better in most suboptimal environmental conditions.<br>

Interestingly, in the large set of functional groups that show low fitness variability we also found many central cellular functions such as those involved in tRNA and amino acid biosynthesis, and central carbon metabolism. It seems likely that the effect of single gene mutations in such functions would be buffered, either by the presence of “redundant” genes, or by less easily identifiable properties of the entire cellular network.<br>

General stress-sensitive groups are posttranscriptionally regulated<br>

If the regulation of the function of a group is mostly on the transcriptional level, the dynamics of transcription should be similar to the dynamics in fitness. If a function is transcriptionally highly variable, but does not lead to variability of fitness, the transcriptional regulation is not relevant for fitness. On the other hand, if a function’s effect on fitness is highly variable but is not mimicked by a variability on the transcriptional level, the effect of the function must be regulated posttranscriptionally. Thus, we can obtain insight into the regulation of a functional group from the relationship between variability of fitness and variability of transcription.<br>

We compared Tvar for transcript data with Tvar for fitness data (Fig. 4B). Clusters a–c were mainly in the upper right segment of the graph, suggesting that these growth related functions are regulated on a transcriptional level, as mentioned above (Supplementary Table 1). In contrast, functional groups in clusters h–j were found in the lower right segment of the graph and thus were transcriptionally less variable than most others. This can indicate that these functions are not regulated on a transcriptional level, or that they are hypersensitive to minor alterations of transcript levels. The latter seems unlikely; besides the fact that there is very little transcriptional variation of the GO-groups in clusters h–j on the level of the individual genes there is no significant correspondence in the direction of change. We checked the direction of transcriptional regulation of the genes in these groups in the DNA damage experiments (Birrell et al., 2001, 2002) and compared the direction of the fitness effect of the deletion mutants in response to the same condition (irrespective of the significance of this alteration), and found no correlation whatsoever. This suggests that
these processes were not observed in the ESR, which is based on transcriptome analyses, because they are not transcriptionally regulated.

Discussion

The relation between fitness and growth

In exponentially growing cells RP genes account for nearly 40% of the total number of polymerase II-mediated transcription initiation events (Warner, 1999). Their transcriptional regulation plays a central role in the control of cell growth through control of translation rate (Schwaller et al., 2004; Wade et al., 2004; Warner, 1999; Warringer et al., 2003). Genes involved in ribosome biogenesis, controlled by Rap1p, Fhl1p, and Sfp1p (Marion et al., 2004; Zhao et al., 2006) are strongly transcriptionally co-regulated, and this transcriptional coregulation also reflected in fitness. Mutations in these genes lead to slow growth (Giaever et al., 1999, 2002; Warringer et al., 2003). In our experiments this is also exemplified by the “sensitivity” of the groups in clusters a–c (Fig. 3) in those experiments that assay growth rate rather than stress adaptation (clusters vii–ix). These mutations also lead to strong resistance to most other environmental perturbations, which can be regarded as external stress (clusters i–vi).

Our method analyzed fitness of all mutants within the population of deletion mutants only rather than in comparison to wild type, which might be one reason for the apparent resistance of ribosome-related deletion mutant groups. However, others have already shown that the ribosome-related mutants are not only the least sensitive within a mutant population, but are indeed stress resistant compared to wild type (Giaever et al., 2002; Tai et al., 2007; Warringer et al., 2003). It has been suggested that the perceived resistance is an artifact of the comparison of stress condition to control condition as deletion mutants are slow growers in optimal conditions, but often display growth rate similar to wild type in stress conditions. It should be borne in mind, however, that the growth rate of these mutants in control conditions is only moderately reduced compared to wild type (Jasnos and Korona, 2007; Warringer et al., 2003), and can be reduced further by additional mutations as predicted by models for genetic interaction (Mani et al., 2008). The mutants could therefore still display condition dependent growth rate reduction comparable to wild-type, which suggests that mutations in these groups lead to actual resistance. Why then do these groups lead to increased fitness in stress conditions? The same set of functional groups are strongly and coordinately downregulated in response to many stresses (Causton et al., 2001; Gasch et al., 2000; Smirnova et al., 2005), and their transcription also strongly correlates with growth rate (Brauer et al., 2008; Castrillo et al., 2007; Regenberg et al., 2006). Growth rate decrease, either induced by mutation of by stress-related downregulation of growth related processes, therefore seems to be important to withstand environmental change. Indeed, in bacterial populations slow growing, genetically identical subpopulations (persisters) have been shown to display significantly enhanced resistance to antibiotics (Balaban et al., 2004). Together, these phenomena suggest a functional role for the environmental stress-induced downregulation of ribosome biogenesis and protein synthesis in growth rate control (Jorgensen et al., 2004; Levy et al., 2007). Such control is key to the successful adaptation to environmental stress.

General requirements for adaptation

Very few transcription factor-regulated mutant groups were significant in the fitness t-profiles. This contrasts sharply to their scores in transcription data. As observed before, transcriptional regulation is only rarely reflected in a growth phenotype (Birrell et al., 2002; Giaever et al., 2002; Haugen et al., 2004; Smith et al., 2006), and many post-transcriptional or network effects disturb the correlation between transcription and fitness. The significant exception from this general observation are the functional groups involved in ribosome biogenesis. Their behavior in fitness experiments correlates highly with their coregulated expression by the transcription factors Fhl1p, Rap1p, and Sfp1p, regulating the transcription of ribosomal proteins and genes involved in ribosome biogenesis (Lee et al., 2002; Zhao et al., 2006).

Comparative analysis of fitness and transcription t-profiles revealed a similar distribution of the numbers of significant GO- and MIPS-based groups in both types of large-scale data (Fig. 2c and d). However, these two types of data represent entirely distinct functional classes. Members of mutant groups involved in common cellular processes often share fitness properties. Fitness of a deletion mutant subjected to a specific condition depends on the fitness contribution of the cellular function in which its gene product participates, and on the adaptive response to that condition. Deletion mutants whose genes are not involved in common processes but are steered by a common transcription factor often lack concerted fitness effects. A possible explanation is that fitness and transcriptome experiments are generally performed on very different time scales; yeast transcriptional changes are studied mostly within the first 2 h of application of stress (one to two generation times), whereas changes in mutant fitness are investigated after 6–16 generations. Indeed, the ESR transcriptional responses are usually transient (Gasch et al., 2000; Gasch and Werner-Washburne, 2002), globally controlled by chromatin modifications (Alejandro-Osorio et al., 2009). These initial responses potentially serve to create a permissive state for more lasting adaptive responses.

Based on global fitness data, it is possible to identify a functional general adaptive response. The two-mode cluster analysis of GO-based fitness t-profiles revealed that the functional groups related to ribosome biogenesis, structure, and assembly perform better in the majority of stress conditions, except for those involved in regulation of growth in response to nutritional signals. Second, we observed hypersensitivity of deletion mutants participating in intracellular transport (vesicle-mediated, endosome, Golgi, and secretion) to most imposed stresses, indicating that these functions may play a central role in response to changes in the environment. Indeed, many genes involved in these functions were identified in two comprehensive chemogenomic screens as multiple drug resistance genes (Dudley et al., 2005; Hillenmeyer et al., 2008). In addition, deletion mutants involved in actin cytoskeleton organization, a function that is also involved in maintaining proper trafficking within the cell, were hypersensitive to many conditions, suggesting a key role for this function in survival and adaptation. A constant (targeted) flux of secretory vesicles, and a balancing flux of endocytic vesicles to recover membrane material are required for cellular remodeling as well as for cell growth (Araujo-Bazan et al., 2008; Higuchi et al., 2009; Kohli et al., 2008; Zonia and Munnik,
2008), also in Saccharomyces (Ayscough, 2005; Moreira et al., 2009; Walther et al., 2006). This appears central for maintenance of cellular homeostasis (Mousley et al., 2008). The high variability of the fitness effects of these mutant groups was not mimicked by a high variability of their transcription. In fact, these groups were among those least variable on the transcriptional level, which suggests that their functions are regulated posttranscriptionally, either by stress-induced regulation of translation (Melamed et al., 2008; Preiss et al., 2003; Shenton et al., 2006; Smirnova et al., 2005), or by post-translational regulation or metabolic activity of the process (Postmus et al., 2008; Wu et al., 2008). Also, we can conclude that the intracellular transport network, or its role in stress adaptation, is not very robust in response to mutations of its individual components. We conclude that the functional groups in clusters h–j, involved in vesicular transport and the actin cytoskeleton, are the key components of yeast’s general adaptive response.

Perspectives for antifungal drug design

The two-mode clustering analysis revealed strong interactions between types of conditions and cellular functions, exemplified by the specific sensitivity of the functions in cluster d to the conditions in cluster iv. Such interactions can be exploited in the development of methodology for fungal growth control in food or medicine. Although other studies have focused on the identification of single protein drug targets, and on the similarity of the effects of compounds (Hillenmeyer et al., 2010; Parsons et al., 2006), our t-profiler analysis indicates cellular functions that can be inhibited to increase the effectiveness of the treatment to inhibit fungal growth (Onyewu et al., 2003; Zakrzewska et al., 2007). Therefore, the two approaches combine to generate approaches for hypothesis driven combinatorial drug design.

Additionally, we observed that mutations in genes with a function in growth rate regulation lead to increased resistance to most chemical stress conditions. This, too, has implications for drug design, as it suggests that conditions, which reduce growth rate, also reduce drug sensitivity. Indeed, perturbation of nutritional signaling renders yeast more susceptible toazole treatments (Robbins et al., 2009). Growth rate reduction may be caused by previous exposure to a drug or treatment, by recent changes, by lack of nutrients, or generally by nonoptimal conditions. Most drug susceptibility testing is mimicked by nonoptimal conditions. Most drug susceptibility testing is mimicked by nonoptimal conditions. Most drug susceptibility testing is mimicked by recent conditions changes, by lack of nutrients, or generally by nonoptimal conditions. Most drug susceptibility testing is mimicked by nonoptimal conditions. Most drug susceptibility testing is mimicked by recent conditions changes, by lack of nutrients, or generally by nonoptimal conditions. Most drug susceptibility testing is mimicked by recent conditions changes, by lack of nutrients, or generally by nonoptimal conditions.

Concluding, we present a simple method that allows for the direct comparison of genome-wide data of various platforms and regulatory levels. The analysis reveals general and specific aspects of yeast adaptation to a changing environment, and provides important insights for the design of anti-fungal treatments.

Acknowledgments

The authors thank Drs. Parsons and Boone for generous sharing of data prior to publication, and Dr. Bussemaker for useful discussions. This research was financially supported by grant APB.5504 from The Netherlands Technology Foundation (STW) (to F.M.K.), and by the BioRange program of The Netherlands Bioinformatics Centre (NBIC), which is supported by The Netherlands Genomics Initiative (NGI) (to J.A.H.).

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References


Address correspondence to:
Dr. Gertien J. Smits
Molecular Biology and Microbial Food Safety
Swammerdam Institute for Life Sciences
University of Amsterdam
Netherlands Institute for Systems Biology
Science Park 904
1098 XH Amsterdam, The Netherlands
E-mail: g.j.smits@uva.nl